Limits to the information gain from lidar measurements

Aniceto Belmonte

Technical University of Catalonia, BarcelonaTech, Dept. of Signal Theory and Communications, 08034 Barcelona, Spain
belmonte@tsc.upc.edu

Received December 3, 2014

Measurements over the return signal are an integral part of lidar remote sensing by which we gather information about the characteristics of specific targets. But how much information is gained by performing a given lidar measurement? By defining Shannon's mutual information of a lidar observation, here we consider the bits of information content on the measurement and describe mathematically the capacity of lidar estimates to represent a corresponding property in the target. For heterodyne Doppler lidars in particular, we have found simple analytical formulas that considers the information gain in mean-frequency estimates.

OCIS Codes: (280.4788) Optical sensing, (280.3640) Lidar, (110.3055) Information theoretical analysis, (010.0010) Atmospheric and oceanic optics, (170.3880) Medical and biological imaging.

The problem of lidar remote sensing is to gain information from the received optical signals about the characteristics of the observed target. In the context of statistical parameter estimation, where a sample of observations depends on a set of target parameters that needs to be estimated, the precision of the signal estimates is commonly investigated as a measure of how well the parameters can be anticipated from the lidar observations. However, the precise amount of information gained by performing a lidar measurement is unknown. Here we use the framework of information theory to provide the limits to the information capacity of a lidar system, or how many bits of information can be obtained from the received signal about the identity of the target that is being observed. Unwanted noise and target speckle effects, those producing great signal variability, limit information gain. Without consideration of the properties of the target to be observed, limits to the information content of lidar estimates in atmospheric, oceanic, or biological media can be quantified.

In this paper we shall reflect on the role of Shannon mutual information in the evaluation of lidar performance and use the analysis of information content of lidar estimates to complement classical studies on precision of signal parameter estimation. Although the work presented in this letter is particular to the lidar remote sensing problem, we note that the subject of information theory with applications to radar has been considered in the past. Early work used information-theoretic ideas to formulate *a posteriori* radar receiver, which in the case of additive white Gaussian noise results in the correlation receiver [1]. Later, the emphasis changed to the problem of radar waveform design for the purpose of extracting information about a target [2]. Radar signature analysis techniques using information theory are also of interest to radar sensor optimization [3].

We consider a lidar receiver that responds to the target-scattering signals of a sequence of N transmitted laser pulses. We denote the lidar responses by a vector $\vec{z}_n = \{z_1, z_2, \ldots, z_M\}$ describing the M stochastic signal samples obtained from the nth lidar pulse. As N of these response vectors are accumulated, then the total lidar response is $\vec{z} = \sum_{n=1}^{N} \vec{z}_n$. The accumulated lidar observation $\vec{z}$ depends on a set of target parameters –such as position, speed, or scattering mechanism– which are usually continuous and need to be estimated. The target parameters are denoted by the scalar variable θ, which have probability $P(\theta)$. The lidar observation vector $\vec{z}$ is also conditional on the transmitted pulse characteristics – such as power, temporal and spectral width, or rate – and other measurement system parameters, which are denoted by σ and are usually known at the receiver. The conditional probability of θ given $\vec{z}$ and σ, written as $P(\theta|\vec{z},\sigma)$, defines the probability that a specific value of parameter θ will occur given transmitted pulse parameters σ and the knowledge that observations $\vec{z}$ have already occurred. If we consider the case of statistically independent lidar pulses, the parameter probability is given by $P(\theta|\vec{z},\sigma) = \prod_{i=1}^{M} P_i(\theta|z_i,\sigma)$, where $P_i(\theta|z_i,\sigma)$ is the probability of a target parameter θ conditional on the ith lidar sample and system parameters σ.

In this paper we focus on the amount of information $I(\theta;\vec{z}|\sigma)$ gained about the target parameter θ by performing the measurement $\vec{z}$ with system parameters σ and consider the Shannon mutual information [4] of the lidar observation. In our case, it is the average amount of information on the target parameter θ that is added by observing the lidar response $\vec{z}$. For a specific parameter θ, the Shannon's mutual information needs to be defined as

$$I(\theta;\vec{z}|\sigma) = h(\theta|\sigma) - h(\theta|\vec{z},\sigma)\,. \tag{1}$$

Here, the term $h(\theta|\sigma)$ is the entropy of the prior distribution $P(\theta|\sigma) = \int d\vec{z}\, P(\theta|\vec{z},\sigma)$ for the target parameter θ and describes the initial ignorance about the parameter:

$$h(\theta|\sigma) = -\int d\theta\;\; P(\theta|\sigma)\; log_2 P(\theta|\sigma)\,. \tag{2}$$

The second term $h(\theta|\vec{z},\sigma)$ in Eq. (1) is the entropy of the posterior distribution $P(\theta|\vec{z},\sigma)$,

$$h(\theta|\vec{z},\sigma) = \int d\theta\; P(\theta)\; \int d\theta\; P(\theta|\vec{z},\sigma)\; log_2 P(\theta|\vec{z},\sigma)\,, \tag{3}$$

and describes the final ignorance about the target parameter θ after the measurements $\vec{z}$. Because the target parameter θ is itself a random quantity, Eq. (3) considers the average over the parameter distribution $P(\theta)$.

In the expression Eq. (1) for the gain of information $I(\theta;\, \vec{z}\,|\sigma)$ from lidar signals – for the purposes of this analysis, we use logarithms to the base two so $I(\theta;\, \vec{z}\,|\sigma)$ is measured in bits, – the entropy of the prior distribution $p(\theta|\sigma)$ for the target parameter θ is considered along with the corresponding entropy of the posterior distribution $P(\theta|\vec{z},\sigma)$ after measurement $\vec{z}$. Entropy is a measure of ignorance, and the gain of information $I(\theta;\vec{z}\,|\sigma)$ is a difference between the initial ignorance and the final ignorance. A lidar system observes a target to decrease the a priori lack of knowledge about that target and, in this sense, the mutual information $I(\theta;\vec{z}\,|\sigma)$ tell us how many bits of information observation $\vec{z}$ provides about the target parameter θ. Lidar measurements provide sufficient information gain to classify parameter θ into a maximum of $\kappa = 2^{I(\theta;\, \vec{z}\,|\sigma)}$ equally likely partitions or choices —by definition, the number of information bits is equal to the logarithm of the number of choices taken to the base two.

At this point we need to consider the characteristics of the posterior entropy $h(\theta|\vec{z},\sigma)$ for lidar observations. As the dimension of the data samples from lidar measurements –given by NM– is usually large, $P(\theta|\vec{z},\sigma)$ will appear to be peaked around its most probable value given observations, i.e., the maximum likelihood estimation of parameter θ. The large sample notion allows the use of a saddle-point method to solve the integral over $P(\theta|\vec{z},\sigma)\, log_2 P(\theta|\vec{z},\sigma)$ in Eq. (3), and approach the final ignorance $h(\theta|\vec{z},\sigma)$ in terms of Fisher's information metric $\mathcal{J}_\theta(\theta)$ of the continuous parameter θ as [5, 6]

$$h(\theta|\vec{z},\sigma) \approx 1/2 \int d\theta\; P(\theta)\; log_2[2\pi e/\mathcal{J}_\theta(\theta)]\,. \qquad (4)$$

Error terms $O(1/NM)$ in Eq. (4) can be neglected for large NM. The well-known Fisher's metric $\mathcal{J}_\theta(\theta)$ –usually considered on the problem of quantifying the efficiency of measuring θ in the lidar responses $\vec{z}$ – relates to the derivative with respect to the target parameter of the lidar response probability. When the lidar measurement of θ considers an efficient maximum likelihood estimator, $\mathcal{J}_\theta(\theta)$ is just the reciprocal of the estimate variance, i.e., $\mathcal{J}_\theta(\theta) = 1/\sigma_\theta^2$ [7].

Finally, putting Eqs. (2) and (4) in Eq. (1) leads to the following expression for the information gain from lidar measurements:

$$\begin{aligned} I(\theta;\vec{z}\,|\sigma) &= -\int d\theta\;\; P(\theta|\sigma)\; log_2 P(\theta|\sigma) \\ &\quad - 1/2 \int d\theta\; P(\theta)\; log_2[2\pi e/\mathcal{J}_\theta(\theta)]\;. \end{aligned} \qquad (5)$$

To research and clarify the implications for lidar remote sensing of Eq. (5), we now turn to a specific case and discuss the information gain from the measurement of heterodyne Doppler lidar return signals [8, 9]. Optical Doppler techniques are of interest in systems operating in the atmosphere, ocean, or biological media. Heterodyne lidars sample the output signals of a complex receiver and vector $\vec{z}$ describes the M complex data samples obtained from the transmitted pulses. Many Doppler lidar estimations are based on the spectrum $\vec{u}$ of the complex data vector $\vec{z}$ and therefore require a sufficiently long data sequence M for the spectrum to be well defined. The vector $\vec{u}$ describe the power levels obtained in the M spectral channels of the spectral sample. Because the spectral vector $\vec{u}$ is the linear discrete Fourier transform of the data vector $\vec{z}$, using the transform vector $\vec{u}$ produces the same information gain than using $\vec{z}$. Note that the lidar return signal spectrum is a convolution of the transmitted pulse spectrum and the target Doppler spectrum. With short pulses, target motion broad barely the signal spectrum and the spectral width of the signal ω will be determined only by the pulse profile and therefore known a priori. For a laser pulse with Gaussian temporal profile, the spectral width ω of the signal is inversely proportional to the pulse temporal width.

The estimations of the mean power – or, equivalently, the signal-to-noise ratio $\hat{\gamma}$– and the Doppler mean frequency shift $\hat{f}$ are important for lidar remote sensing. Although both of these parameters have unbiased, efficient maximum likelihood estimators, here we will be concerned only with $\hat{f}$ Doppler estimates providing information about the radial component of the target velocity.

If we assume that SNR $\hat{\gamma}$ and signal spectral width ω are known, the performance of the maximum likelihood estimator of $\hat{f}$ is well known [10-13]. Earlier work on the processing of radar and lidar signals found analytical formulas describing asymptotic bounds to the performance $\mathcal{J}_{\hat{f}}(\hat{f})$ of the mean frequency estimator [10-11]. Exact, numerically efficient representations of the Fisher's metric $\mathcal{J}_{\hat{f}}(\hat{f})$ have been presented in detail [12]. Approximate $\mathcal{J}_{\hat{f}}(\hat{f})$ expressions for the estimation of the mean frequency, which can be easily computed in most practical situations, are also relevant on bounding Doppler lidar performance [13]. Without lost of generality, in our analysis we will consider these simplified expressions for $\mathcal{J}_{\hat{f}}(\hat{f})$.

In order to solve the information gain Eq. (5) for the mean frequency $\hat{f}$, we need to express the prior and posterior entropies, $h(\hat{f}|\sigma)$ and $h(\hat{f}|\vec{z},\sigma)$, respectively. For the posterior probability, as the Fisher's metric $\mathcal{J}_{\hat{f}}(\hat{f})$ is independent of the mean frequency $\hat{f}$, i.e., $\mathcal{J}_{\hat{f}}(\hat{f}) = \mathcal{J}_{\hat{f}}$ [13], the entropy of the posterior distribution Eq. (4) simplifies to $h(\hat{f}|\vec{z})$=$1/2\;\; log_2(\,2\pi e\;/\;\mathcal{J}_{\hat{f}})$. For the prior probability, as the pulse spectral width ω determines unambiguously the signal spectral width, the probability of $\hat{f}$ is conditional on the uncertainty ω of the received signal frequency. Also, as the lidar measurement uses a frequency estimator, the uncertainty of the estimator σ_f needs to be added to the prior probability. Under the usual assumptions of Gaussian pulse profile and efficient, normally distributed maximum likelihood frequency estimator, the best prior distribution $P(\hat{f}|\sigma)$ for the mean frequency $\hat{f}$ is a Gaussian with variance $\omega^2 + 1/\mathcal{J}_{\hat{f}}$. On these terms, the prior entropy Eq. (2) results $h(\hat{f}|\,\sigma) = 1/2\; log_2(\,2\pi e\,(\omega^2 + 1/\mathcal{J}_{\hat{f}})\,)$.

Now, using the prior $h(\hat{f}|\sigma)$ and posterior $h(\hat{f}|\vec{z},\sigma)$ entropies above, the information gain $I(\hat{f};\vec{z}\,|\sigma)$ for the mean frequency $\hat{f}$ can be stated using Eq. (5) as

$$I(\hat{f};\vec{z}\,|\sigma) = 1/2\ \ log_2\big(\,1 + \ \omega^2\,\mathcal{J}_{\hat{f}}\,\big)\,. \qquad (6)$$

This result offers an intuitively appealing explanation. If the estimation problem has small variance, which for an efficient estimator is given by the reciprocal of the Fisher information $\sigma_f^2 = 1/\mathcal{J}_{\hat{f}}$ [7], then the occurrence of this measurement brings much information. If $\omega^2\mathcal{J}_{\hat{f}} = \omega^2/\sigma_f^2 \gg 1$, then $I(\hat{f};\vec{z}\,|\sigma) \approx log_2(\,\omega/\sigma_f\,)$ the information gain is logarithmic in the ratio of pulse bandwidth ω to estimation uncertainty σ_f. This ratio ω/σ_f defines the maximum number κ of equiprobable partitions of $\hat{f}$. When the ratio ω/σ_f is small, the information gain $I(\hat{f};\vec{z}\,|\sigma) \approx 1/2\,(\omega^2/\sigma_f^2)\,log_2 e$ is also small and the number of likely partitions of $\hat{f}$ tends asymptotically to $\kappa = 1$. Note that the information gain $I(\hat{f};\vec{z}\,|\sigma)$ from the parameter measurement cannot be less than zero as the ignorance about the parameter after the measurement cannot be larger than the initial ignorance.

We need to consider briefly the information capacity of the mean-frequency measurement. In essence, information capacity can be used to establish fundamental limits to the problem of lidar information measurement that is independent of any particular approach to solving it. It provides an upper bound on lidar information gain that can be approached but not exceeded. The capacity is defined as the maximized mutual information over all the priors, i.e., $C = \max_{P(\hat{f})} I(\hat{f};\vec{z}\,|\sigma)$ [4]. For a given mean and variance, Gaussian distributions always reach this maximum. Consequently, by considering a Gaussian laser pulse profile and an efficient Gaussian estimator, the result presented in Eq. (6) achieves capacity and defines an upper bound to Doppler information gain. This is an important result as the assumptions of a Gaussian pulse profile or an efficient Gaussian estimator may not be well founded in some practical situations. Even in these cases where $I(\hat{f};\vec{z}\,|\sigma)$ cannot achieve capacity, Eq. (6) is useful describing the upper bound to the information gain of mean frequency $\hat{f}$ estimates.

The information gain in mean-frequency estimates, calculated from Eq. (6), is shown in Fig. 1 as a function of different system parameters. Typical parameters for a heterodyne lidar are chosen to be signal bandwidth normalized to the sampling frequency $\omega = 0.1$, and number of complex samples per estimate $M = 128$. The plots consider different values of lidar pulse return accumulation (spectral accumulation) N. The signal energy per mean-frequency estimate is measured with the effective number of coherently detected photons $\hat{\gamma}$ for the accumulated return data sample. For a photon noise limited heterodyne receiver, $\hat{\gamma}$ corresponds to the ratio over the entire spectral bandwidth of the signal power to noise power, or wideband SNR per observation time.

Information gain $I(\hat{f};\vec{z}\,|\sigma)$ as a function of SNR $\hat{\gamma}$ is shown in Fig. 1 (a). The figure considers when a discretely sampled spectral is measured from only a lidar pulse return ($N = 1$), and when several of these sample spectra are accumulated to define the total sample spectrum (N of 10, 100, and 1000). For small SNR, when the accumulated signal is too poor to produce good estimates, by gradually increasing the accumulated energy guarantees a steady gain of information. But, at large SNR, when performing the lidar measurement has eliminated most ignorance about the mean frequency $\hat{f}$, the amount of information gain reaches a threshold. Information content in further received signals from the same target mostly echo already known information.

These two different behaviors of the information gain $I(\hat{f},\vec{z}\,|\sigma)$ can be easily quantified by considering the analytical formulas describing asymptotic bounds to Fisher's metric $\mathcal{J}_{\hat{f}}$ of the mean frequency estimator [11]. On the one hand, for small SNR $\hat{\gamma}$, Fisher's metric $\mathcal{J}_{\hat{f}} \approx NM\hat{\gamma}^2/4\pi^{1/2}\omega^3$ and the mutual information Eq. (6) simplifies to

$$I(\hat{f};\vec{z}\,|\sigma) = 1/2\ log_2\big(\,1 + \ NM\hat{\gamma}^2/4\pi^{1/2}\omega\,\big)\,. \qquad (7)$$

This result explains the logarithmic increases in information with SNR $\hat{\gamma}$ shown in Fig. 1 (a). On the other hand, for large SNR, $\mathcal{J}_{\hat{f}} \approx NM/12\omega^4$ and information gain reach a threshold independent of SNR $\hat{\gamma}$:

$$I(\hat{f};\vec{z}\,|\sigma) = 1/2\ \ log_2(\,1 + \ NM/12\omega^2\,)\,. \qquad (8)$$

Equation (8) describes how the information threshold for high SNR grows logarithmically with accumulation order N.

To illustrate these information asymptotic bounds, Fig. 1 (a) considers information gain for a discretely sampled spectral measured from a set of $N = 1000$ lidar pulse returns, an otherwise usual accumulation order in practical heterodyne lidars. The dashed, broken curve shows asymptotic bounds for this accumulation level. In this case, information gain shows values as high as 9 bits, corresponding to a maximum $\kappa = 512$ number of equiprobable partitions for $\hat{f}$. Note that for this accumulation number, information gain increases rapidly with SNR but only when the amount of received energy reach a minimum level. The minimum SNR required to reach a mere 2-bit gain of information is a considerable high 10-dB per accumulated sample spectrum.

As the accumulation order N increases, and the signal spectrum become less noisy, it is easy to understand how spectral accumulation produces a higher information gain from lidar estimates. To assess the effects of accumulation from a different perspective, Fig. 1 (b) shows the information gain at several fixed levels of $\hat{\gamma}$ (SNR of 0, 10, 20, and 40 dB) as the accumulation order N increases. After reaching a maximum at a certain accumulation order, the information gain falls and ultimately goes to zero. This is not surprising, as the estimation of the mean frequency involves obtaining the location of the spectral peak at the signal spectra in the presence of random noise, and positioning the precise peak is easier when the spectral feature of the signal is not buried in the noise. Usually, when the accumulated SNR $\hat{\gamma}$ is large –and the energy into one single pulse is enough to produce optimal signals for make good estimates, – increasing the accumulation N of independent samples result in a decrease of the statistical uncertainty and, consequently, an increase of the information gain. However, when the accumulated SNR $\hat{\gamma}$ is small, and the energy received from a single pulse averages less than 1 photon per

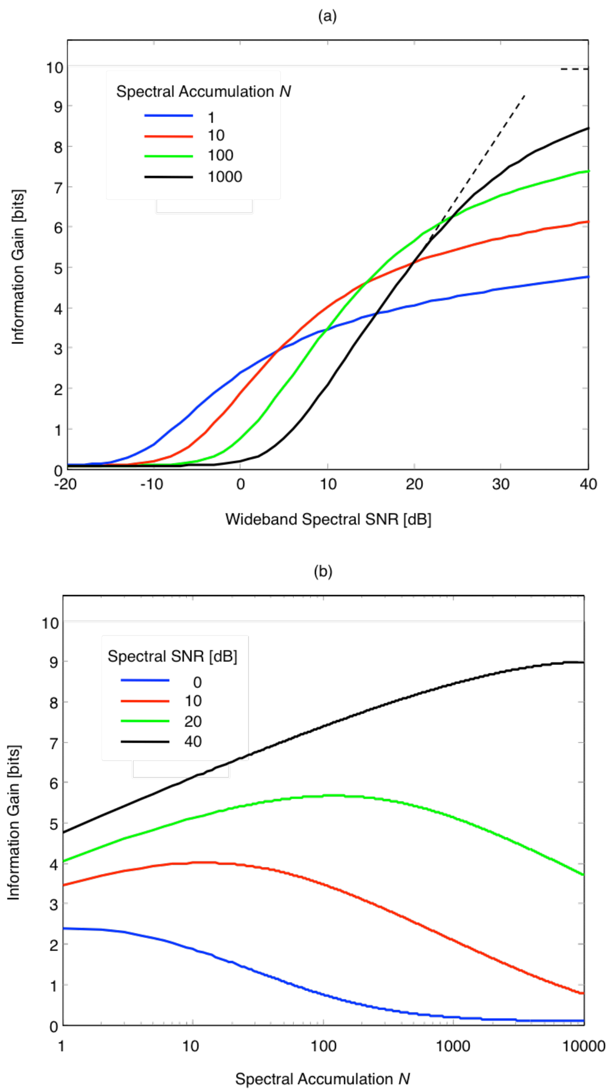

Fig. 1. For heterodyne Doppler lidars, the curves represent bits of information gain $I(\hat{f};\vec{z}\,|\sigma)$ obtained by mean-frequency estimators $\hat{f}$ over the signal returns $\vec{z}$ about the Doppler shift of the observed target. The curves consider a signal bandwidth, normalized to the sampling frequency, of $\omega = 0.1$ and use $M = 128$ spectral channels. (a) The results are a function of the coherent SNR (number of photons $\hat{\gamma}$ coherently detected) per accumulated sample spectrum and ponder different values of lidar spectral accumulation N. (b) Now the curves are a function of the spectral accumulation N and consider different coherent SNR levels per accumulated spectrum ($\hat{\gamma}$ of 0, 10, 20, and 40 dB).

sample, spectral signals are too noisy to produce good estimates. In this case, using more energetic pulses, to concentrate all the energy into a few sampled spectra, is the best strategy to overcome noise. Information gain increases with spectral accumulation –and eventually reaches a maximum, – when every sampled spectrum collects enough energy to prevail upon noise. As expected, the maximum value of information gain increases logarithmically with the SNR $\hat{\gamma}$ per observation time.

In conclusion, the problem of finding the information content of a lidar observation has been confronted. We identify Shannon's mutual information of a lidar measurement as the bits of information gained by performing it, independent of the target state on which it is performed. The current work depicts in Eq. (5) a simple dependence of the mutual information on the precision of signal estimates that can be used to explain the basic information capacity of lidar measurements.

We study new logarithmic measures of information in lidar sensing. For the use of the heterodyne receiver scheme in particular, we have analyzed with Eq. (6) the information gain in mean-frequency estimates and found how many bits of information can be extracted from heterodyne signals as a function of different measurement parameters (see Fig. 1). We have shown that Eq. (6) achieves capacity and describes an upper bound to the information content of Doppler estimates.

In this letter we have considered how much information we can infer from a particular set of lidar data about the observed target in terms of mutual information. As a necessary complement to conventional estimation theory, and to capture all important features of signal information content, the subject of this work has been the study of novel logarithmic measures of information in lidar observations. The rigorous framework of the information theoretic approach may offer new perspectives on the field of lidar remote sensing.

This research was partially funded by the Spanish Department of Science and Innovation MICINN Grant No. TEC 2012-34799.

REFERENCES

1. P. M. Woodward, I. L. Davies, *Proc. IEE* **99**, 37-44 (1952).
2. M. R. Bell, *IEEE Trans. Inform. Theory* **39**, 1578-1597 (1993).
3. J. A. Malas, K. M. Pasala, Proc. *IEEE* Radar Conf. **2008**, 26-30 (2008).
4. T. M. Cover, J. A. Thomas, *Elements of Information Theory* (Wiley-Interscience, 2006).
5. J. Rissanen, *IEEE Trans. Inform. Theory* **42**, 40-47 (1996).
6. N. Brunel and J.P. Nadal, MIT Press Neur. Comp. **10**, 1731-1757 (1998).
7. L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis* (Prentice Hall, 1991).
8. T. Fujii and T. Fukuchi (Eds.), *Laser Remote Sensing* (CRC Press, 2005).
9. Claus Weitkamp (Ed.), *Lidar: Range-Resolved Optical Remote Sensing of the Atmosphere* (Springer, 2005).
10. M. J. Levin, *IEEE Trans. Inform. Theory* **11**, 100-107 (1965).
11. D. S. Zrnic, *IEEE Trans. Geosci. Electron.* **17**, 113-128 (1979).
12. R. Frehlich, *IEEE Trans. Geosci. Remote Sensing* **31**, 1123-1131 (1993).
13. B. J. Rye and R. M. Hardesty, *IEEE Trans. Geosci. Remote Sensing* **31**, 16-27 (1993).